© Universiti Tun Hussein Onn Malaysia Publisher's Office

IJIE

http://penerbit.uthm.edu.my/ojs/index.php/ijie
ISSN : 2229-838X e-ISSN : 2600-7916

The International Journal of Integrated Engineering

Dorsal Hand Vein Identification using Transfer Learning from AlexNet

Waheed Ali Laghari[1], Kim Gaik Tay[1*], Audrey Huong[1], Yann Yee Choy[2], Chang Choon Chew[1]

[1]Department of Electronic Engineering, Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, Johor, 86400, MALAYSIA

[2]Department of Mathematics and Statistics, Faculty of Applied Science and Technology, Universiti Tun Hussein Onn Malaysia, Johor, 86400, MALAYSIA

*Corresponding Author

DOI: https://doi.org/10.30880/ijie.2022.14.03.012
Received 26 December 2021; Accepted 04 June 2022; Available online 20 June 2022

Abstract: Dorsal hand vein pattern is a highly secured biometric system that has been used in many applications due to its non-contact attributes. Prior studies focused on investigation of different deep networks for hand vein classification task using different training parameters. It is the aim of this study to propose the use of systematic fine-tuning system for identifying the best parameters value for enhanced model learning efficiency. In this study, pre-trained AlexNet was trained using Bosphorus hand vein database for identification of 100 users. The experiments were carried out using original images, and preprocessed (cropped) images for comparison. The testing accuracies of these datasets were compared following tuning of training parameters, namely training and testing split ratio, number of epochs, mini-batch size and initial learning rate. It was observed that the testing accuracy of the model trained using cropped images is higher than that using the original images. The model from preprocessed dataset shows a good testing accuracy of 96 % using a split ratio of 90:10, epoch 50, mini-batch-size of 10 and an initial learning rate of 0.0001. The performance of our trained model is more superior than many reported results in the field. In future, the performance of this classification system may be further enhanced with automatic search of parameters for improved model training efficiency.

Keywords: Hand vein, Bosphorus, AlexNet, training parameter, transfer learning

1. Introduction

Nowadays, an automatic identity recognition system based on biometric methods plays a vital role in many applications, such as for use in high-security environments, public banks, education sectors, etc [1]. In other words, it is the classification of human characteristics, either physiological or behavioral [2, 3]. Studies in verification of individual's physical characteristics, such as fingerprints [4], iris of the eye [5], palm vein [3] and dorsal hand vein [6], have gained an increase interest because it cannot be shared, duplicated and easily lost [7]. In contrarily, behavioral characteristics such as keystroke [8] and voice [9] can be easily cracked and fooled.

Although the fingerprint authentication is our society's standard identity recognition system, it can be easily stolen and duplicated by capturing and printing the pattern on gelatin material board [6]. Similarly, iris recognition system is sensitive to variations of light, and it cannot be implemented with sunglasses or glasses on. Other problems associated with conventional face recognition systems are processing speed and storage, image size and quality, surveillance angle, light variations, and inter-class variability [10].

*Corresponding author: tay@uthm.edu.my

A dorsal hand vein pattern is a capillaries network of blood vessels that carry deoxygenated blood from the body to the heart beneath a person's skin. Hand vein pattern identification was first proposed for use as a biometric marker in 1992 [11]. Following that, many researchers diverted towards the dorsal hand vein biometric system because, unlike the iris, its pattern can be seen with naked eyes. Unlike fingerprints, vein pattern is hidden under the skin, so it is difficult to be forged [12]. In addition, this system is contactless, hygiene and is not affected by the external conditions, such as temperature and humidity change. Everyone possesses a different pattern of veins, even among twins. Moreover, the pattern remains unchanged during the aging process. It was concluded in [7] that the dorsal hand vein recognition is a unique, reliable and consistent biometric system.

Neural Networks (NN) is an increasingly popular method used in classification

of data. It can be classified into different categories, which include Artificial Neural Network (ANN), Recurrent Neural Network (RNN), Convolution Neural Network (CNN), and Deep Neural Network (DNN). The recognition accuracy depends on the learning capability of the model in extracting and recognition importance features of dataset used in the training.

Unlike dorsal vein authentication using manual ANN [13-14] machine learning methods that require manual feature extraction (such as statistics and Gray-Level Co-Occurrence Matrix (GLCM) features [13] and local binary pattern feature [14]), CNN can automatically extract features, and at the same time, reduce the size of features map for multi-label categorization tasks. A CNN has minimal image pre-processing steps because it combines image segmentation, feature extraction, and classification in one system [7]. Designing CNN from scratch requires a lot of times to identify the weight of the deep network and huge datasets to train for a better accuracy. Hence transfer learning plays an important role because it uses pre-trained CNN models that were originally trained on images from the same domain. There are several pre-trained CNN models available, such as AlexNet [15], GoogLeNet [16], DenseNet [17], ResNet [18], VGGNet [19], SqueezeNet [20], etc. The classification performance of these models depends on the hyperparameters settings used in the network training.

Previous work in [21] segmented Near Infrared (NIR) images of dorsal hand vein to recognize Region of Interest (ROI), before they are enhanced through CLAHE (Contrast Limited Adaptive Histogram Equalization) and Gaussian filter. Following that, pretrained AlexNet, Reference-CaffeNet, VGG16, and VGG19 were trained using the collected dataset to classify new hand vein images. The training accuracy of AlexNet, VGG16, VGG19, and Reference-CaffeNet are reported as 99.1 %, 99.61 %, 99.7 %, and 99.33 %, respectively

Another work in [7] developed a palm dorsal hand vein recognition system using a CNN. They used Badawi dataset in [22], which contains 5 right-hand images per 50 subjects, and split the dataset with a ratio of 80 %/ 20 % for training and testing purpose. This dataset was trained on two CNN models (i.e., 6-32-50 and 6-32-100 feature maps sizes in four convolution layers) and achieved identification accuracy of 98.8 % and 97.6 %, respectively.

Recent work in [23] demonstrated two classification approaches using Badawi [22] and Bosphorus dorsal hand vein datasets [24]. In the first approach, they applied pre-trained CNN models (AlexNet, VGG16, and VGG19) for automatic features extraction features and Error-Correcting Output Codes (ECOC) with Support Vector Machine (SVM) and K-Nearest Neighbor (K-NN) algorithms for classification. The second one is using transfer learning of the above-mentioned pre-trained CNN models for automatic features extraction and classification. They used different split ratios of 80:20, 70:30, and 60:40 for recognition of the dorsal hand vein in the first approach. The Bosphorus dataset produces an accuracy of 95.51 % using transfer learning from AlexNet with an epoch of 50, mini-batch size of 10 and initial learning rate of 0.0001.

Kumar [25] captured the dorsal hand vein images using market available NIR camera (VF620). They obtained 99.60 %, 98.46 % and 97.99 % accuracies for good, medium and low-quality images, respectively, using VGG Net-16 and a split ratio 50:50 for training and testing processes.

Even though transfer learning from AlexNet has been applied on Bosphorus dorsal hand vein dataset [24] by researchers [23] using 1500 original images without pre-processing and varying number of epochs (10, 20, 30, 40,50), mini-batch size of 10 and initial learning rate of 0.0001. It is our intention to demonstrate the improvement in these results using our adopted image cropping procedure and propose the use of systematic and orderly system for fine-tuning of training hyperparameters namely training and testing split ratio, number of epochs, mini-batch size and learning rate.

2. Methodology

Fig. 1 shows the proposed image processing and classification flow. The process started with image acquisition, followed by image labeling according to the individual class. The original images and images with ROI manually cropped from the dataset were stored and used for further processes. These images were resized using *imresize* function in MATLAB according to the input requirement of AlexNet. Training parameters were determined by varying split ratio of training and testing, maximum epochs, mini batch size and initial learning rate. The detailed description is provided in the following subsections.

2.1 Image Acquisition

The employed images were acquired from Bosphorus dorsal hand vein dataset [24]. This dataset was collected using NIR technology consisting of a monochrome NIR camera with an infrared lens [26]. These three-channel images are of gray level, have 8-bit depth and a resolution of 300×240 pixels and are stored in BMP format. The dorsal hand dataset contains 1575 left and right-hand images acquired under different experiment conditions. In this paper, 1500 images were selected from this dataset categorized into 100 classes. Each class contains 15 images as shown in table 1. Fig. 2 shows an example of a original right-hand vein image from Bosphorus dorsal hand vein dataset [24].

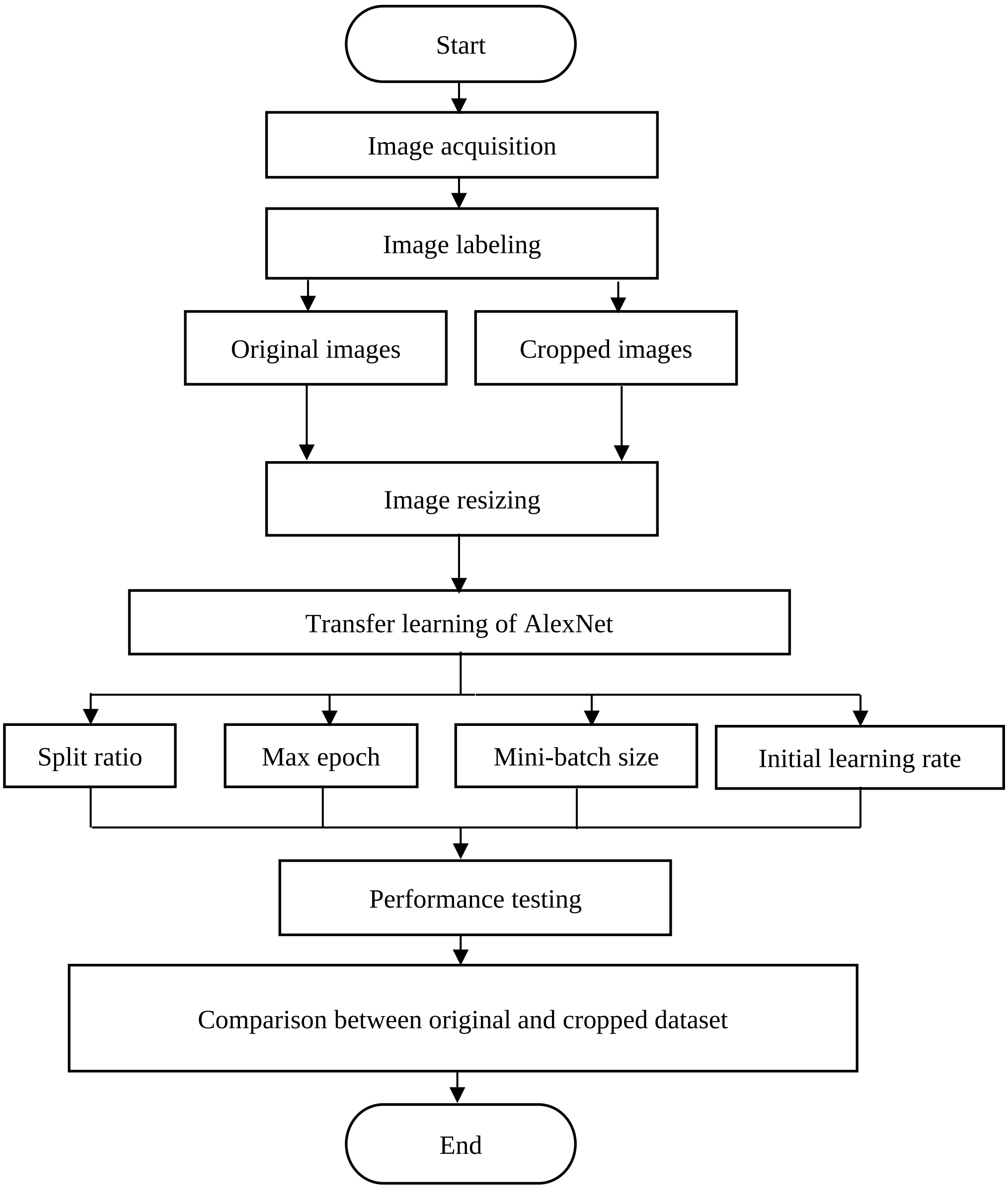

Fig. 1 - Flow chart of the proposed classification flow

Table 1 - Number of dorsal hand vein images collected under different experiment conditions in each class

No of Images	Hand sidedness	Condition
3	Left	Normal/ at rest
3	Left	After carrying a bag for one minute
3	Left	After squishing an elastic ball (opening and closing) for one minute
3	Left	After placing ice piece at the back of the hand
3	Right	Normal/ at rest

Fig. 2 - An example of right-hand image under at rest condition

2.2 Image cropping

A total of 1500 images (corresponded to 100 users) of the Bosphorus dorsal hand vein dataset were called into MATLAB platform and manually cropped using the *imcrop* function. The blue rectangular portion around a dorsal hand vein image shown in Fig. 3 for demonstration was chosen to extract the ROI required for further process. The output of this cropping process is shown in Fig. 4.

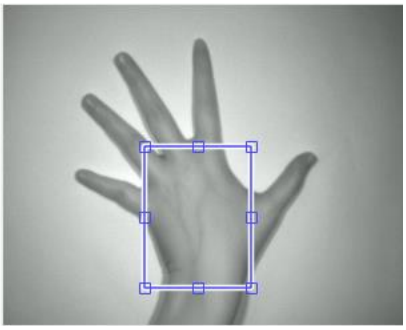

Fig. 3 - The identified ROI of a dorsal hand vein image

Fig. 4 - The cropped image

2.3 Image resizing

The original and cropped images were resized to 227×227 using *imresize* function without filtering and enhancement processes. These images were applied consecutively to train AlexNet for dorsal hand vein recognition task.

2.4 Pre-trained Alexnet and model training

AlexNet was first proposed by Alex Krizhevesky and it was the winning architecture of the ImageNet contest held in 2012. It is a deep CNN model used for visual recognition and classification of the dataset. AlexNet was chosen because it has shallower network as compared to other CNN models, thus less computation time and memory are needed. Transfer learning technique was adopted as we have limited number of images each class (i.e., 15 images). Training using transfer learning can transfer the weight learned from the pre-trained CNN models (previously trained on large number of images) to solve the limited data size problem in this study.

The training simulation was run on Intel(R) Core (TM) M-5Y71 CPU @ 1.40GHz with 8GB RAM. The original and cropped images were input into AlexNet. The performances of cropped and original images were analyzed and compared for various training parameters (number of epochs, mini-batch sizes, initial learning rates and split ratio of training and testing) as listed below:

- Epoch: 10, 20, 30, 40 and 50
- Mini-batch size: 10, 32, 64 and 128
- Initial learning rate: 0.0001, 0.001, 0.01 and 0.1
- Split ratio (training and testing): 90:10, 80:20, 70:30, 60,40 and 50:50

3. Results and Analysis

The effect of training parameters (number of epochs, initial learning rate and mini-batch, and split ratio of training and testing) on testing accuracy are presented in the following. We adopted a systematic and well-organized means in identifying the best hyperparameters set.

3.1 Effect of split ratio on testing accuracy

This fine tuning begins with tuning split ratio. For demonstration purposes, Stochastic Gradient Descent with momentum (SGDM) was used for training. Table 2 shows the original and cropped testing accuracies when using various split ratios (training and testing) while other training parameters were fixed using transfer learning from AlexNet. Cropped images give higher accuracy than the original images using different split ratios for training and testing). This result shows that the higher the number of training images, the better its testing accuracy. Overall, cropped images achieve an average increase in testing accuracy of 3 % as compared to the original dataset.

Table 2 - The testing accuracy of original and cropped images using a different split ratio from AlexNet

Split Ratio	Testing Accuracy (%)		Training Parameters
	Original	Cropped	
90:10	91	96	Solver (SGDM) Initial learning rate (0.00010) Max Epoch (50) Mini-batch size (10)
80:20	87.7	89.3	
70:30	87.5	90.75	
60:40	83.9	86.83	
50:50	81.2	84	

3.2 Effect of epoch on testing accuracy

Following the results in Table 2, the split ratio of 90:10 was used for the following process. This subsection investigates changes in testing accuracies with the number of epochs varied. Fig. 5 shows the testing accuracies of model trained with original and cropped images using different number of epochs. The highest testing accuracy (96%) with epochs 50 and using the cropped images. This figure shows that the superiority in the testing performance of cropped images as compared to the original images except for epoch of 10.

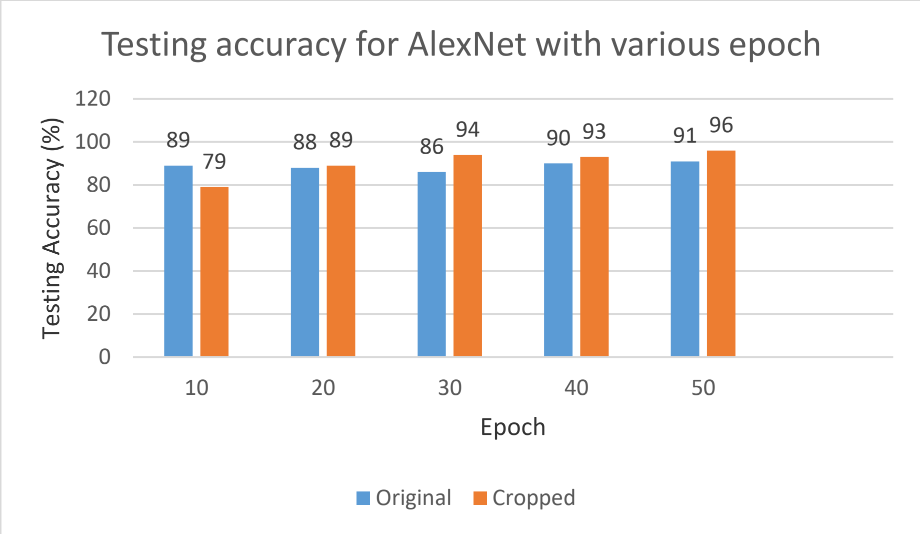

Fig. 5 - Testing accuracies of original and cropped images with different epoch number. Other fixed parameters include a split ratio of 90:10, mini-batch size of 10, initial learning rate of 0.0001.

3.3 Effect of mini-batch size on testing accuracy

Next, based on the results in Table 2 and Fig. 5, a split ratio of 90:10 and the number of epochs = 50 were used in the following process with the mini-batch size varied. The initial training rate follows Table 2. Similar to the earlier findings, fig. 6 shows model trained using cropped images gives higher testing accuracy than that using the original images with different mini-batch sizes. The highest accuracy was achieved with a mini-batch size of 10. On the contrary, original images achieve the highest accuracy of 92% with a mini-batch size of 32.

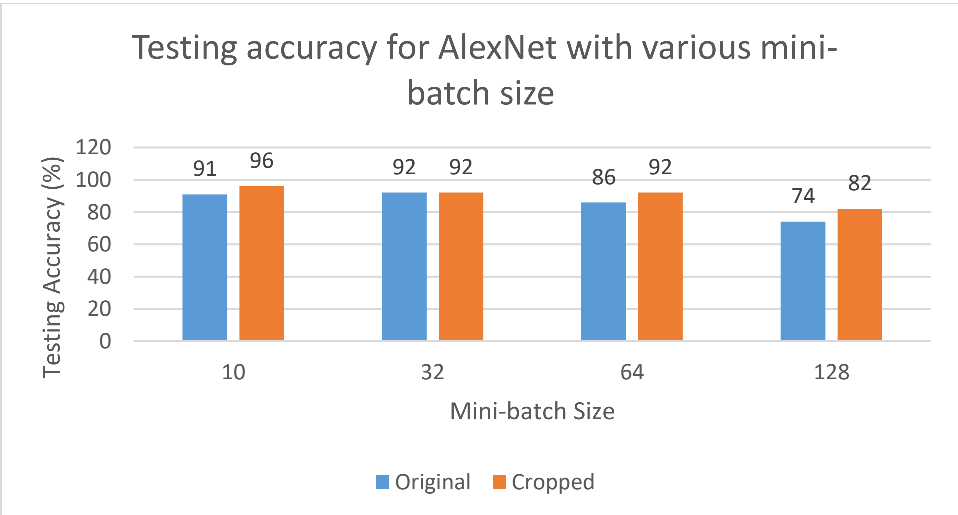

Fig. 6 - Testing accuracy of model trained using original and cropped images with different mini-batch sizes, a split ratio of 90:10, number of epochs = 50, learning rate of 0.0001.

3.4 Effect of initial learning rate on testing accuracy

Using the best results from Table 2, Figures 5-6 (i.e., split ratio of 90:10, number of epochs = 50 and mini-batch size of 10), initial learning rate was varied in this subsection. Fig. 7 reveals that the cropped images achieve higher accuracy than the original images with various initial learning rates. This trend is consistent with our earlier findings. It is observed that as the initial learning rate increases, the testing accuracy decreases. The initial learning rate of 0.01 and 0.1 produces zero testing accuracy. Overall, the model trained with cropped images produced the highest testing accuracy (96%) with an initial learning rate of 0.0001.

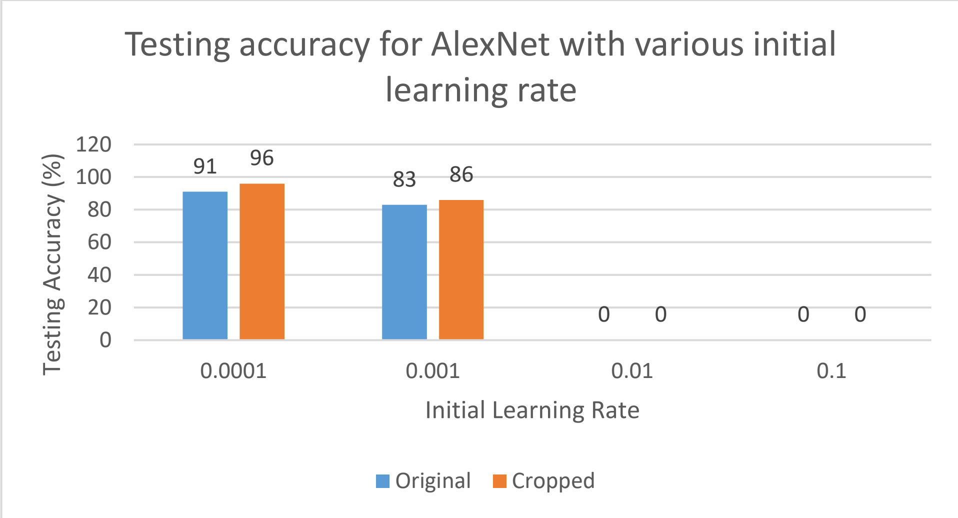

Fig. 7 - Testing accuracy of model trained with original and cropped images using various initial learning rates, with a split ratio of 90:10, number of epochs = 50, and mini-batch size of 10.

4. Discussions

This study demonstrates a systematic and organized way of fine-tuning parameters (i.e., split ratio, epoch number, mini-batch size and initial learning rate) important in improving model learning efficiency. This work was carried out on AlexNet model transfer-learning using original and preprocessed (cropped) dorsal hand vein images dataset. Based on the results shown in Table 2 and Figures 5-7, it was observed that the model trained using cropped images that contain only a particular portion of an image produced a higher accuracy for dorsal hand vein recognition when compared to the original images. The reason behind this is likely because of the removal of backgrounds and fingers, and hence enhancement of data representation. This suggests preprocessing via image cropping is necessary for improved classification of dorsal hand vein images. We found the highest accuracy of 96% using the best training parameters set obtained thus far, i.e., an epoch of 50, mini-batch size 10, an initial learning rate of 0.0001, and a split ratio (training and testing) of 90:10.

It was found in Table 2 that the higher the number of training images, the better the testing accuracy. In general, training dataset is used for a CNN to extract important features, and establish a correlation between these features and class label, while testing dataset are used to check the performance of the trained network. Therefore appropriate selection of split ratio is important to guarantee convergence of the model to the training data [27]. The increased in training data allows the model to learn different features and their variability. This leads to improvement in the classification performance.

Meanwhile our results in Fig. 5 shows a positive relationship between the epoch number and testing accuracy. The improper selection of epoch number may lead to two major problems namely overfitting and underfitting. Overfitting is the mixture of training data and noise (irrelevant information), which negatively affects the performance and hinders the accuracy of the model. Underfitting (often associated with small epoch) refers to the situation when a model failed to learn well on the training dataset causing poor classification performance [28]. From our findings in Fig. 5, it may be suggested that epoch 20 is the cut-off value for the employed model to learn the useful features using the cropped image dataset, and to prevent underfitting problem. On the contrary, the testing accuracy decreases with an increase in the mini-batch size in Fig. 6. This is likely because large number of mini-batch size would take too long to achieve convergence (reduced learning ability) when compared to the small mini-batch size. The latter produces significant regularization effect and low generalization error [29].

Similarly, the increase in initial learning rate leads to a decrease in the testing accuracy in Fig. 7. This study observed zero testing accuracy using initial learning rate of 0.01 and 0.1. We attributed this to the infeasibility of the model to learn important features under these learning rates. In all cases, SGDM was used for model training the dataset. This learner produces comparable model performance as Adaptive Moment Estimation (Adam) and Root Mean Square Propagation (RMSProp) using a shorter computing time [30]. Our findings (for cropped dataset) are better than many reported works in [7] and [23], but we lose out on classification performance to [25]. This is likely due to the high image quality used in [25]. In the future, this work may be expanded to explore the use of automatic fine-tuning system for optimization of training parameters to improve performance of this classification system.

5. Conclusion

In conclusion, model trained using cropped images produced higher accuracy for dorsal hand vein recognition as compared to that using the original images. This is because cropped images contain only a particular portion of the dorsal hand vein rather than the whole image, hence enhancing representation of the data. We demonstrated the systematic way of identifying the best hyperparameter set for model training. The cropped images achieve the highest testing accuracy of 96 % using an epoch of 50, mini-batch size 10, an initial learning rate of 0.0001, and a split ratio (training and testing) of 90:10. This result is better than many existing works in the field, suggesting its feasibility for field application using this sustainable and resource-efficient recognition system.

Acknowledgement

The authors would like to thank the Ministry of Higher Education Malaysia for financially supporting the work under the Fundamental Research Grant Scheme for Research Acculturation of Early Career Researchers (FRGS-RACER) scheme RACER/1/2019/ICT02/UTHM//1 and registrar Universiti Tun Hussein Onn Malaysia.

References

[1] Singaram, M., Praveena, P., Sabitha, J., & Shalini, M. (2019). Dorsal hand vein authentication system. Int. J. Innov. Res. Sci. Eng. Technol, 8, 2566–2571.
[2] Jain, A. K., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. IEEE Trans. Circuits Syst. Video Technol, 14, 4–20.
[3] Lin, C. L., & Fan, K. C. (2004). Biometric verification using thermal images of palm-dorsa vein patterns. IEEE Trans. Circuits Syst. Video Technol, 14, 199–213.
[4] Faridah, Y., Nasir, H., Kushsairy, A. K., Safie, S. I., Khan, S., & Gunawan, T. S. (2016). Fingerprint biometric systems. Trends Bioinforma, 9, 52–58.
[5] Alice Nithya, A., & Lakshmi, C. (2015). Iris recognition techniques: a literature survey. Int. J. Appl. Eng. Res, 10, 32525–32546.
[6] Heenaye, M. M. N. M., and Subramanian, R. K. (2009). A study of dorsal vein pattern for biometric security. Univ. Mauritius Res. J, 15, 17–25.
[7] Rajalakshmi, M., Ganapathy, V., & Rengaraj, R. (2017). Palm-dorsal vein pattern authentication using Convoluted Neural Network (CNN). Int. J. Pure Appl. Math, 116, 525–532.
[8] Ali, M. L., Monaco, J. V., Tappert, C. C., & Qiu, M. (2017). Keystroke biometric systems for user authentication. J. Signal Process. Syst, 86, 175–190.
[9] Singh, N., Agrawal, A., & Khan, R. A. (2018). Voice biometric: a technology for voice based authentication. Adv. Sci. Eng. Med, 10, 754–759.
[10] Zulfiqar, M., Syed, F., Khan, M. J., & Khurshid, K. (2019). Deep face recognition for biometric authentication. 1st Int. Conf. Electr. Commun. Comput. Eng. ICECCE, 1-6.
[11] Sontakke, B.M., Humbe, V. T., & Yannawar, P. L. (2018). Automatic ROI extraction and vein pattern imaging of dorsal hand vein images. International Journal of Scientific Research Engineering & Technology (IJSRET), 4, 1678-1683.
[12] Gupta, P., and Gupta, P. (2014). Extraction of true palm-dorsa veins for human authentication. ACM Int. Conf. Proceeding Ser, 14, 1–8.
[13] Chin, S. W., Tay, K. G., Huong, A., & Chew, C. C. (2020). Dorsal hand vein pattern recognition using statistical features and artificial neural networks. IEEE Student Conference on Research and Development (SCOReD), 217-221.
[14] Chin, S. W., Tay, K. G., Chew, C. C., Huong, A., & Rahim, R. A. (2021). Dorsal hand vein authentication system using artificial neural network. Indonesian J. Elec. Eng. Comp. Sci, 21, 1837-1846.
[15] Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2017). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60, 84-90.
[16] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolution. Proc. 2015 IEEE Conf. on Computer Vision and Pattern Recognition, 1-9.
[17] Huang, G., Liu, Z., Maaten, L. V. D., & Weinberger, K. Q. (2017). Densely connected convolutional networks. Proc. 2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2261-2269.
[18] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 770-778.
[19] Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. Proc. 3rd Int. Conf. Learn. Represent ICLR, 1-14.
[20] Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2017). Squeezenet: Alexnet-level accuracy with 50x fewer parameters and< 0.5 mb model size. Proc. 5th Int. Conf. Learn. Represent ICLR, 1-13.

[21] Wan, H., Chen, L., Song, H., & Yang, J. (2017). Dorsal hand vein recognition based on convolutional neural networks. Proc. - 2017 IEEE Int. Conf. Bioinforma. Biomed. BIBM, 1215–1221.
[22] Shahin, M., Badawi, A., & Rasmy, M. (2010). Multimodal biometric system based on near-infra-red dorsal hand geometry and fingerprints for single and whole hands. World Academy of Science, Engineering and Technology, 4, 268-283.
[23] Al-Johania, N. A., & Elrefaei, L. A. (2019). Dorsal hand vein recognition by convolutional neural networks: feature learning and transfer learning approaches. Int. J. Intell. Eng. Syst, 12, 178-191.
[24] DotNetNuke Corporation. Bosphorus hand database. http://bosphorus.ee.boun.edu.tr/hand/Home.aspx.
[25] Kumar, R., Singh, R. C., & Kant, S. (2021). Dorsal hand vein-biometric recognition using convolution neural network. Adv. Intell. Syst. Comput, 1165, 1087–1107.
[26] Belean, B., Streza, M., Crisan, S., & Emerich, S. (2017). Dorsal hand vein pattern analysis and neural networks for biometric authentication. Stud. Informatics Control, 26, 305–314.
[27] Nguyen, Q. H., Ly, H. B., Ho, L.S., Al-Ansari, N., Le, H. V., Tran, V. Q., Prakash, I. & Pham, B. T. (2021). Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. Math. Probl. Eng., 2021, 1-15.
[28] Afaq, S., & Rao, S. (2020). Significance of Epochs on training a neural network. Int. J. Sci. Technol. Res., 19, 485–488.
[29] Kandel, I., & Castelli, M. (2020). The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. ICT Express, 6, 312–315.
[30] Dogo, E. M., Afolabi, O. J., Nwulu, N. I., Twala, B., & Aigbavboa, C. O. (2018). A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks. 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), 92-99.